Don't move!

ANNABEL CORMACK & NEIL SMITH

Abstract

We consider head-movement, A′-movement, and A-movement. We argue that of these three kinds of apparent movement, only head-movement exhibits displacement in any real sense, and that all three kinds of structure are made available directly by Merge. Merge itself is no longer implicitly restricted to function argument application, but is enriched to include other combinatorial operations derived from Combinatory Categorial Grammar. We argue that the re-interpretation of 'movement' is not only possible, but desirable on empirical and conceptual grounds. We further consider why natural language should exhibit these apparent displacements.

1 Introduction

The Minimalist program of Chomsky (1995) sets an agenda of reducing the properties of grammar to those with external motivation, thus explaining away the apparent 'imperfections' of NL. The initial emphasis was on properties determined by 'legibility' conditions set by the interfaces to thought/action and articulation/perception. Other legitimate sources of constraints on NL that Chomsky has appealed to are processing considerations (the search space determined by well-formedness conditions) and memory cost (passing partial information to the interfaces to lower memory load). There has been no recent appeal to constraints which might arise from learnability considerations.

One of the areas which Chomsky has considered is 'movement' or 'displacement'. In a recent paper (Chomsky 2001), he suggests that Move be reinterpreted as 'Internal Merge', and hence that the endeavour to motivate movement is otiose. We do not see that the notion of 'internal merge' is a natural one, nor that its semantic and syntactic properties emerge without stipulation. Further, we will argue that both standard A-bar movement and A-movement are properly seen as formed entirely by (external) merge of independently formed items, but that it is nevertheless reasonable to ask why the grammar contains the 'traces' and syntactico-semantic operators that make these apparent displacements possible.

2 Grammars and combinators

Early formal grammars relied on phrase structure rules and lexical entries to licence the merger of words and phrases. A PS rule such as 'VP → V NP' licensed a VP [*eat peas*], provided that *eat* was entered in the lexicon with the contextual information '— NP', and *peas* as NP. Principles and Parameters style grammar partly eliminated the redundancy by replacing the PS rules with the general X′ schema, and the lexical specifications were partially replaced by s-selection or theta grids. However, these devices relate only to lexical heads and their complements, so that additional devices relating to adjuncts, specifiers, and functional heads had to be introduced. The relation of the structures generated by syntax to those needed for semantics was apparently accidental: the syntactic structure (after movement and deletion of various kinds) was fed to the cognitive interpretive interface. In earlier models this was not done until the whole sentence was complete. Experimental work has shown that in listening or reading, real time processing involving inference is carried out during, rather than just at the end of, the input of a whole sentence (see e.g. Frazier, 1988); yet the competence model of grammar proposed offers no obvious way of accounting for this possibility. In the Chomsky (1999, 2000) model, the smallest unit available for interpretation is the Phase (excluding its head and specifiers). The smallest phase is *v*P, but in a simple sentence, this will not in fact be accessible until the whole sentence is complete, so it is of no obvious use in real time input processing.

Combinatory Categorial Grammars offer a simple solution to all these problems (but standardly, at a cost which we discuss below). First, as in Montague grammar and HPSG, syntax and semantics go hand in hand, with parallel syntactic and semantic selection and result given for every lexical item and for every instance of merge. Second, 'composing' combinators allow a merger of non-standard constituents, so that as an alternative to a right-branching structure like (1a), we may have the left-branching (1b).

(1) a [Johnny [must [not [leave]]]]]

b [[[[[Johnny] must] not] leave]

In (1a), although the individual items *Johnny*, *must*, and *not* are successively accessible as the sentence is heard, no larger constituents are available for inference until we get to *leave*. Under the mergers of (1b), successive constituents [*Johnny must*], [*Johnny must not*], and so on, are made available for on line inference.[1]

The semantic selection is that given under the assumption that every item has a meaning which can be characterised by its contribution to a truth value.[2] Thus the meaning for some propositional complement verb simplistically requires the values for two entities, represented by DP and CP, before it can give a saturated V projection, which would be a phrase with a truth value. This is coded by giving it a category V[/D/C], with the parallel semantic type <(t/e)/t> (for convenience, the brackets in such representations may be omitted; and we are ignoring intensionality). In addition, CCG is explicit about the semantics of the Merge operation itself. Two words or phrases when merged do not just sit inertly, either syntactically or semantically. A procedure of some sort determines the result of the merge, giving both the syntactic category and the semantics of the mother (and some phonological representation as well, though much of the time we will omit this). Implicit in early Phrase Structure grammars is the assumption that the only available form of combination is function argument application: a function is saturated by applying it to its arguments; a lexical head is likewise saturated syntactically and semantically by applying it to its arguments. If *think* has the category V[/D/C], and meaning $\lambda p \lambda x[(\text{think}'.x).p]$ with type <(t/e)/t>,[3] then when merged by function argument application with [*Mary sneezed*], a saturated projection of C of type <t>, the output will be of category V[/D] and meaning $\lambda x[(\text{think}'.x).[\text{Mary sneezed}]']$ of type <t/e>. This is the proper category, meaning and type for the verb phrase [*think Mary sneezed*], which still requires one further argument.

CCG, however, allows more than one way of putting two items together under Merge, so these need to be specified. The combinator responsible for function argument application is <u>**A**</u>, which is a functor taking two operands.[4] The lexical entry for <u>**A**</u> licenses structures of the sort shown in (2), where the infix dot is function argument application, and the symbol '_' is used for concatenation:

---

[1] See Sperber & Wilson 1995: 205 ff.

[2] As in Cormack 1989, this is taken to be a syntactic representation of truth with respect to the model of the world held by the speaker.

[3] We reverse the standard type notation so that it runs in parallel with the syntactic categories.

[4] For the combinator <u>**A**</u>, see Szabolcsi (1992: 249). For the standard Combinatory Categorial Grammar version of composition using <u>**B**</u>, see Steedman (e.g. Steedman 1993, 2000).

(2) semantics: **A** *f* *x* = *f* . *x*
categories: **A** F/X X = F
types: **A** α/β β = α
phonology: **A** *a* *b* = *a_b*[5]

The problem of left to right interpretation and on-line inference is solved by introducing a composing combinator **B**. This combinator is defined in (3). At its simplest, it re-brackets a string, leaving the final meaning intact.

(3) semantics: (**B** *f* *g*) *x* = *f* . (*g*. *x*) i.e. **B** *f g* = $\lambda x\, f(g(x))$
categories: **B** F/G G/X = F/X
types: **B** α/β β/γ = α/γ
phonology: **B** *f* *g* = *f_g*

Other family members, such as $\mathbf{B}^2$ shown in (4), are often notated with a superscript, but are here given the same symbol **B**.

(4) $((\,(\mathbf{B}^2\, f\ g).\, x)\ .\, y) \ = f\, .\ ((g\ .\, x)\ .\, y)$

In a CCG with re-bracketing or composing combinators of this sort, the result of every merge can be sent to the interfaces for interpretation. The examples in (1) will be parsed as in (5), with every constituent, including the non-standard ones in (5b), given an appropriate meaning and category.

(5) a [**A** Johnny [**A** must [**A** not leave]]]

b [**A** [**B** [**B** Johnny must] not] leave]

In standard CCG, word order variation both within and between languages is accounted for by assuming that there are directional variants of the operations above, encoded with either directional slashes or order-reversing combinators. There is no direct analogue of head movement, so that lexical categories have to bear the burden of

5 See Steedman 2000 for a combinatorial account of intonational phonology.

accounting for the relevant data. This leads to apparently wide variation between languages, so that as with most surface structure grammars, it is hard to discern what belongs to UG, what to parametrisation, and what to lexical idiosyncrasy. For example, with a verb like *help*, CCG offers differing categorial selection for English and German subordinate clauses as shown in (6) and (7):[6]

(6) English: (VP/VP)/NP λyλPλx help′.(Py).y.x type: <t/e/(t/e)/e>

[$_{VP}$ [$_{VP/VP}$ [$_{VP/VP/NP}$ helps] [$_{NP}$ John]] [VP to build a house]]

(7) German: (VP\NP)\VP λyλPλx help′.(Py).y.x type: <t/e/(t/e)/e>

[$_{VP}$ [$_{VP/VP}$ [$_{NP}$ dem Hans] [[VP ein Haus zu bauen] [$_{VP\backslash VP\backslash NP}$ hilft]]]]

These differ not only in the direction in which the V heading the VP expects to find its internal arguments, but in the order in which the different arguments are selected. C-selection and s-selection are not in correspondence in the case of the German.

Despite our claim that movement is unnecessary, our aim is to preserve the insights into UG offered by P&P style grammars with movement. Several things follow with respect to our use of CCG. The order should be taken care of instead by the assumption that if arguments appear on the 'wrong' side of a head, it is because there has been "head-movement". We make the restrictive assumption that syntactic and semantic selection correspond, so that learning the lexicon is simplified, and syntactic command gives logical scope in simple cases, as is standard in P&P grammar. The German c-selection order shown is thus ruled out. We also wish to preserve the insight that all semantic selection by non-binding heads is for simple (saturated) categories such as <e> for 'entity' and <t> for 'proposition'.[7] Consequently, we must eliminate the selection for <t/e> in the type for HELP given above: see section 5 for a solution.

We assume there are mental processes corresponding to combinators, and that combinators are represented mentally. Thus they may enter in to linguistic

[6] These are adapted versions of those in Steedman (2000: 139).

[7] We necessarily assume that a binding determiner is syntactically as well as semantically a two-place operator: it selects not only for a projection N/D (i.e. an NP) but also for X/D (e.g. VP or some other unsaturated projection; see Cormack 1998).

representations. The present paper is an exploration and justification of 'movement' under this assumption.

The assumption that **A** and the other combinators are linguistic entities is a departure from standard practice. We assume that the combinators are present in syntactic structure; they can then act as a locus of head-movement, as in Cormack 1998. We take it that the arity (valency) of each element is fixed (with adjuncts selecting for their hosts), and that functors precede their arguments at LF. This means that no brackets are actually needed in the LF representation, nor in the definition of merge, since we argue that merge obtains LF representations directly (see Chomsky 2000: 133-4). Our assumption that the combinators are uniformly head-initial gives a fixed word-order,[8] so that so far these moves apparently give us no opportunity to account for word order variation in NL. However, standard CCG offers combinatory accounts of A′-movement: we exploit these and sketch our variants of them in section 5. In section 6, we introduce a treatment of A-movement which depends on a combinator dubbed **R** (for 'Raising'). Our most radical departure from the current conceptions of CCG arises from our treatment of head-movement, which we discuss in section 3. This departs as well from the current Minimalist versions of head-movement. Sections 4 and 7 discuss other kinds of 'displacement' that may arise within the framework.

3 Merge and the minimalist program: head movement

The Minimalist program has eliminated D-structure and S-structure. After heads are merged into a structure, movement may apply either to Formal Features or, if required by morphophonological conditions, to a head or phrase. Where head-movement takes place, it is as a result of morphophonological necessity. Conceptually, morphophonological conditions apply to the PF-interpretable part of a linguistic sign, so that the question arises as to whether head-movement has any semantic (LF-interpretable) effect. Our hypothesis is that the answer to the question is 'No', as follows from the axiom in (8):

(8) The LF-part of a head is merged in the position in which it is LF-interpreted

[8] This provides an alternative to Kayne's (1994) Linear Correspondence Axiom. The problem raised by type lifting must be treated: we have proposals in a paper in preparation.

Adopting (8) as an axiom reduces the options available to the syntactician. Equally, and desirably, on natural assumptions about L1 acquisition, it reduces the options available to the child learner. Once the relative semantic scope of the various heads is known, the relative order in which they are merged at LF is known, and vice versa.

More controversially, we apply the same reasoning to the phonological representations of lexical entries as to their meanings. On a lexicalist hypothesis, the lexicon may contain morphologically complex phonological items. We see no reason to suppose that these are Merged anywhere other than where they are heard. Minimalist considerations then dictate that the PF part of a sign is merged where it is heard. If there is apparent movement, this is because the sign is SPLIT, and the PF-interpretable part of the sign is displaced relative to the LF-interpretable part, as shown in the axiom in (9).

(9) The PF-part associated with a head is merged where it is PF-interpreted

Where can PF-parts appear? We claim that contra LFG and other theories, there is no independent generative device accounting for the PF-ordering of PF-parts of lexical items (for LFG's 'constituent-structure', given by a PSG, see e.g. Dalrymple et al. 1994: 211, Bresnan 2000: chapter 6).[9] Only LF-parts systematically have selection features, and it is these which determine structure, and in our view, order too. Thus we postulate that the position of PF-parts is parasitic on that of LF-parts: PF-parts must appear at positions constructed by LF-parts.

For example, consider the morphophonological word *could* in (10), which relates to the LF interpretable CAN and PAST.[10]

(10) The dog could not reach the ball.

---

[9] Jackendoff (1997) has three generative devices. The PF device is seen by Chomsky (2000) as taking the output of Spell-Out and interpreting it further (so what we call the PF-interpretable part of a sign is just that: it can be interpreted by the independent PF component). The syntactic device in fact gives the PF-ordering of words; there is a separate semantic device because of Jackendoff's espousal of a semantic representation which is decompositional HPSG has again a syntactic device on the one hand and instructions for assembling meanings on the other. CCG generates PF-interpretable surface structures, and in some instantiations (e.g. Steedman 1996) has a separate predicate-argument structure comparable to LF.

[10] This is an example simplified for expository purposes. In Cormack & Smith 2000a, we give more detailed examples, and argue that both V and Aux in English may have PF forms relating also to $Infl^T$, rather than directly to T.

PF (temporal) order: *could not*

LF (scope) order: PAST [NOT [CAN […

Small caps are used for the LF-interpretable parts of signs (i.e. for meanings, more or less), and italics for the PF-interpretable items. Note that because the LFs PAST and CAN are not adjacent in the LF representation, it cannot be the case that *could* has a single amalgamated meaning. Instead, *could* occurs in the lexicon minimally as in (11), with a simple PF and a set of (two) LFs:

(11) PF: $_{\{\text{Aux, T}\}}$ *could* LF: { $_{\text{Aux}}$CAN, $_{\text{Tense}}$PAST}

Since CAN and PAST are merged at different places in the tree, with Tense having scope over the modal, at least one of the two LF-heads must be 'Split' from its related PF-part. If *could* is heard at the LF-position of PAST, then there is no point assuming that it is inserted at the position of CAN, rather than PAST. We therefore assume that it actually is inserted at the position of PAST, and hence that the sign corresponding to the lexical item 'can' is split into LF-related and PF related parts which are merged in different places, as shown in (12). The relation of PF-items and associated LF items is established in a derivation by checking.

(12) a LF: [$_{\text{T}}$ PAST [$_{\text{Pol}}$ NOT [$_{\text{Aux}}$ CAN …

PF: [$_{\{\text{T, Aux}\}}$ *could* [$_{\text{Pol}}$ *not* [$_{\text{Aux}}$ …

The two dotted lines indicate where a checking relation is needed. We envisage that this is accomplished by percolation and feature-selection, and that the outputs are subject to a set of soft constraints of the type familiar in OT. One of the constraints, requiring that we split as few signs as possible (equivalent to STAY in OT, Grimshaw 1997: 374), ensures that the PF *could* is merged under either LF PAST or LF CAN. Another constraint favours the PF part of a split sign being merged higher than the related LF part: this dictates that the PF *could* is merged at LF PAST rather than at LF CAN.

The idea, then, is that where the lexicon provides a single PF-interpretable item corresponding to a complex of more that one LF-interpretable item, then the PF-interpretable item will usually appear at the position of the highest of the related LF-interpretable items. It follows that there is no need for head-movement in the usual sense, and that there are no intermediate or indeed final head-traces, nor a Copy Theory of head-movement. However, contra Chomsky 1999: 31, we claim that the displacement of heads is part of narrow syntax rather than belonging within a phonological component. Notice that our Split Signs proposal avoids the well-known problems relating to the alternative head-adjunction model (Stump 1996: 237), but we owe an explanation of the Mirror Principle (Baker 1985). We discuss some of the issues briefly below.

In a Split Sign system, the burden of constructing PF forms corresponding to more than one LF head must be borne by the lexicon. Regular morphology is the product of morphophonological rules which we take to inhere in the lexicon. Irregular forms are more highly specified, and will override defaults. We note that there are two extreme scenarios relating to a range of functional projections inflectionally related to a head. The members of the set may occur only in a fixed order, or in a variable order. We assert without proof that in the fixed-order case, it is possible to devise a recursive set of morphophonological rules which will induce Mirror Principle orderings.

For the freely ordered case, a Mirror Principle solution is clearly impossible, because our checking hypothesis requires that the categorial content of the PF be an unordered set. We predict then that there are languages where the Mirror principle is flouted. Hyman (2001) argues that Bantu provides examples (see also Hyman & Mchombo 1992).

Bantu languages typically have a series of morphemes that may be suffixed to the verb, designating Causative, Applicative (e.g. intrumental, shown in the informal LF below as **use**), Passive, and Reciprocal. Suppose we take these to correspond to heads at LF, where such heads may occur in various scope orders. It is usually suggested that the occurrence of the corresponding suffixes satisfies Baker's Mirror Principle. In other cases, as in the Chichewa in (13), Hyman argues that the morpheme ordering is constrained by adherence to a pan-Bantu template –**V**–C–A–R–P– (but note that overall the situation in Chichewa is much more complex, and includes Mirror Principle effects). The (a) forms are as given by Hyman; the (b) forms give an informal syntactically expanded meaning. FV is a 'final vowel'.

(13) a	abúsa	a-ku-thámáng-its-ir-a	ndodo	mbúú
	shepherds	3sg-prog-**run-cause-app**-FV	sticks	goats

'the shepherds are making the goats run with sticks' b [the shepherds [**use** sticks [**cause** [the goats **run**]]]]

(14) a alenjé a-ku-tákás-its-ir-a akází mthíko
hunters 3sg-prog-**stir-cause-app-**FV women spoon

'the hunters are making the women stir with a spoon'

b [the hunters [**cause** [the women **use** a spoon [**stir**]]]]

In our theory, this contrast arises from the assumption that at least the V and Applicative heads must be Split Signs, with related lexical entries partially as in (15).

(15) PF: $_{\{V,\ Appl,\ Caus,\ \ldots\}}$ stem$_X$-*its-ir-*

LF: $\{_V X,\ _{Caus}$ CAUSATIVE, $_{Appl}$ APPLICATIVE, ...$\}$

This entry does not determine in what order the LF-parts appear, so that in particular, it is used in both (13) and (14), inducing the 'template' rather than 'mirror' ordering of the applicative and causative morphemes. Hyman cites Chimwi:ni (Mwiini G. 40) as being strictly templatic in this sense; only the CARP order of morphemes is possible whatever the meaning (Hyman refers to Abasheikh 1978: 28, which we have not seen).

We note that it seems to be impossible on the Split Sign analysis, at least without stipulation, to have a morphology which allows for free ordering of LF heads and simultaneously satisfies the Mirror Principle.[11] Thus we cannot give a split signs account of the Quechua data shown in the examples in (16a) and (17a) (from Baker 1985; originally from Muysken 1981:296):

(16) a Maqa-naku-ya-chi-n
beat-recip-dur-**cause-**3S

'He$_j$ is causing them$_i$ to beat each other$_i$'

b He **cause** [them [**RECIP beat**]]

[11] Checking as proposed in Chomsky 1995: 195 to obtain Mirror Principle effects is essentially stipulative.

(17) a Maqa-chi-naku-rka-n

beat-cause-recip-pl-3S

'They$_i$ let someone$_j$ beat each other$_i$'

b They [**RECIP** [**let** [someone **beat**]]

What we might deduce from this is that some languages do not accrue the verb-related meanings by projecting functional heads, but rather by constructing complex morphological meanings in the lexicon. In relation to Chichewa, Alsina (1999) argues on empirical grounds for just such a lexical account. However, it is possible to construct a syntactic version, provided that the set of Infl heads and the verb are contiguous at D-structure. Given the informal meanings represented in (16b) and (17b), this may seem impossible. However, using composing combinators can contrive this, as we show in section 7 for V-selecting heads combined with **R***.[12]

Note however that if complex meanings are constructed in this way, certain meanings available in a Split Sign or head movement version are no longer available. For instance, in contrast to what is possible in a Split Sign or head-movement account, there can be no skipped adverbial or negation meanings such as we find in (12); if passive is included in a complex verbal morpheme, there can be no phrasal passive (as in [*the shirt was* PASS [*ironed dry*]]); and since in serial verb structures a single instance of T has local scope over two verbs, there can be no serial verb constructions in a language where a verb in the immediate scope of T must form a complex with T at LF. For this reason, we must reject Alsina's claim that all morphologically realised argument-changing operations (such as passive and causative) are operations within the lexicon.

4 Split Signs and the displacement of phrases

In Cormack & Smith 2000b, we argue that the Split Sign theory of head movement extends readily to certain instances of what would standardly be seen as phrasal movement, specifically, one kind of "fronting" in English. We showed that there were good reasons to distinguish this displacement from that of standard A′ movement. In

[12] Whether in syntax or the lexicon, we would suppose that the heads are combined using composing combinators, to account for the inheritance and alterations to argument structure.

general, under a Bare Phrase Structure analysis (Chomsky 1995a), and under CCG assumptions, there should be few differences between heads and phrases. If the two parts of a unitary sign such as CAN may be split before merge, then so can a phrase, in principle.

Consider (18), in the context of showing off birthday presents:

(18) This tie,	**Fred**	brought		(bold type is used to mark the main stress)
H LH	H*	L%		(notation of Pierrehumbert 1980)[13]

In order to account for the displacement of the phrase *this tie* in (18), we postulate a semantically vacuous head F**on**, c-selecting for TP. F**on** makes no overt phonological contribution, and must by virtue of its associated PF entry in the lexicon, host PF material from elsewhere. The PF part of some word or phrase within TP must in either case be merged at F**on** to satisfy this PF requirement. The kind of material is determined by its PF properties. Following and adapting Ackema & Neeleman (2000 a, b), we take it that if an m-selecting phonologically empty head has a morphological shape, even in the form of a null affix, then it can only host a head. If however it is phonologically radically null, then the head may host a phrase. The morphological instructions for building the PF lexical entries for a morphologically null and radically null "affixal" head X with host Y are given in (19) and (20).

(19) Head X is morphologically null: PF{X,Y} = PF-Y + ∅

(20) Head X is radically null: PF{X,Y} = PF-Y

Here is a simple example of a F**on** structure. For clarity, we simplify as far as we can, and use notation such as XP and specifiers without commitment to their validity. The sentence in (18) is constructed by merging the LF-interpretable pieces shown in (21a), meeting the selection requirements of the various heads.

(21) a LF: $[_{F\ \mathbf{on}}$ IDENTITY$[_{TP}$ FRED $[_{T}$ PAST $[_{VP}$ $[_{V}$ BRING $[_{DNP}$ THIS TIE]]]]]]

b PF: $[_{F\ \mathbf{on}}$ this tie $[_{TP}$ Fred $[_{T}$ *e* $[_{VP}$ $[_{V}$ brought $[_{DNP}$ *e*]]]]]]

[13] We use this notation for convenience only; we are not committed to all aspects of Pierrehumbert's theory, according to which the representation in (18), for instance, would be ill-formed.

So far as semantics is concerned, the LF-part corresponding to the fronted phrase is in the clause-internal non-fronted position. In other words, it will behave interpretatively as if it is 'reconstructed' to the internal position.

We argued that the use of a head such as F **on** would be motivated by pragmatic, rather than semantic or syntactic considerations. Consequently, it gives the appearance of allowing 'optional movement'.

5 Merge and the minimalist program: A′-movement

In earlier papers, we have discussed our non-movement accounts of 'A′-movement'. We illustrate with a second analysis of "fronting" in English. Consider a simple focus structure, as in (22), used in answer to the question, *Which tie did Tom bring?*

(22) **This tie**, Tom brought
H* L H L H%

There are good reasons to suppose that the LF structure for such a sentence must have something giving the effect of a quantifier-variable interpretation, as is standard for A′-movement. In particular, the fronted material may or may not be interpreted as if reconstructed, as we see in the two interpretations of (24) below. We will however argue that even when the interpretation requires reconstruction, the fronted phrase is in fact merged and interpreted in the initial pre-subject position.

Our second "fronting" mechanism involves a Case-licensing head **Gap**, which selects for TP. Following Cormack (1999) and the assumptions of CCG, we assume that the mode of discharge of theta-roles of lexical heads is mediated locally by a Combinator, where the choice of Combinator is determined by the syntactic Case licensing available. If a head assigns [+Case] then a theta-role has to be discharged by some suitable phrase, using the combinator **A**. Thus if the sentence is well-formed, the presence of **Gap** entails that some theta-role is available for discharge, and that it is discharged by some phrase. That is, **Gap** requires both the insertion into the structure of the fronted 'argument' phrase, and the presence of a 'gap' in the TP (given by a trace), so that some theta-role is available for the argument to discharge.

Consider then the structure (23) postulated for (22). We assume, still following Cormack (1999), that a noun-phrase is a higher order phrase selecting to the right for a predicate which itself selects for a D, and hence that the noun phrase precedes the phrase providing the theta role which it discharges. Any discrepancy between the LF order and the PF order is to be accounted for by Split Signs.

(23)

LF: $[_{GapP}$ $[_{DNP}$ THIS TIE] $[_{\mathbf{Gap}}$ *ID* $[_{TP}$ TOM $[_{T}$ *PAST* $[_{VP}$ $[_{V}$ BRING] $[_{DNP}$ *TRACE*]]]]]]

PF: $[_{GapP}$ $[_{DNP}$ this tie] $[_{\mathbf{Gap}}$ *e* $[_{TP}$ Tom $[_{T}$ brought $[_{VP}$ $[_{V}$ *e*] $[_{DNP}$ *e*]]]]]]

We assume that traces are distinct grammatical objects, with their own syntactic, semantic, and phonological properties. Under this interpretation, traces are not excluded by the Inclusiveness principle of the Minimalist programme (Chomsky 1995: 228). In order to account for the scope ambiguity in examples like (24), where the fronted noun phrase may have scope either above or below *intend*, we postulate a lower type and a higher type trace.

(24) Two chapters, John intends to read *t*/*T*

We argued in Cormack & Smith (2000b) for this particular interpretation of traces on the basis of the necessity to account for the licensing of fronted NPIs, where we must be able to characterise a command condition as well as a scope condition. Under this hypothesis, in (23) the relation between the LF THIS TIE, and the chosen trace is not mediated by movement. Note however that this interpretation of traces is only viable to the extent that a phrase including a trace is a well-formed syntactic object.

The syntactic properties and meaning of the TP can be produced compositionally by using the 'function composition' combinator <u>**B**</u>, instead of the ordinary function-argument combinator <u>**A**</u> to merge items. Essentially, <u>**B**</u> passes an undischarged argument selection up the tree. This will yield the two versions of (24) for the two distinct trace values of the D-headed complement of the verb. The combinatorial structure will be partially as in (25):

(25) [<u>**A**</u> [two chapters] [<u>**B**</u> [<u>**B**</u> [<u>**B**</u> John intends] to read] *t*/*T*]]

The **Gap** structure, as exhibited for instance in (23), makes available at LF the gapped TP, a phrase which is effectively predicated of the fronted phrase. The interpretation of the fronted phrase as providing a contrastive or exhaustive focus is the natural result. This is because the interpretation of focus typically requires just such a separation into a predicate and its argument, where the argument is the focussed phrase and the predicate is the 'background' (see Jackendoff (1972: 245-247), Chomsky (1977: 91-92), Krifka (1991), Rooth (1992) and others). It has also been argued by various authors (e.g. Brody 1990: 209) that a (single interrogative) *wh*-phrase is always focussed. We take it then that '*wh*-movement' is also parasitic on **Gap**, so that there is no movement involved here either.

6 Merge and the minimalist program: A-movement

By assumption, the structures delivered by LF for further processing are interpretable in the Language of Thought (LoT) (Fodor 1975, Sperber & Wilson 1995). The minimalist program requires the default assumption that the interpretation be direct. If for instance we posit variables in LF, then we must posit variables in LoT; if we propose the Copy Theory of movement, where the lowest copy is 'interpreted' as a trace, then again we must posit variables in LoT. We are of course accustomed to variables in the predicate calculus, but there are good reasons to suppose they are not part of LoT. Fortunately, it turns out that they are not required in NL either.[14]

(26) [likely [*t* to fail]]

Fragments such as that in (26) are clearly interpretable at both interfaces; the LF interpretation does not fail because there is an unbound variable at the position marked by a trace. We suggest that UG allows the constituents to be combined directly. However, in the standard version of a P&P type grammar, the only available mode of combination for the discharge of lexical selections is implicitly function argument application. But here, the required meaning cannot be obtained in this way, since *likely* expects a complete proposition as its internal argument. Rather than introducing lambda operators into the interpretation to bind the trace, we postulate a combinator **R** (for 'raising') that gives directly the correct meaning and category for the combination of phrases required, as in Cormack 1999. It does this not in terms of argument positions (because the external argument for (26) for instance is not supplied), but in terms of semantic and syntactic selections. Note that some such semantic operator must in any case be available to give the semantics of the essentially syntactic solutions to raising

[14] For a treatment of bound-variable pronouns without variables, see Jacobson 1999.

and control suggested by transmitting theta roles (Cormack 1989, Hornstein 1998, 1999, 2001), or by postulating action at a distance (Manzini & Roussou's (2000) variant of Chomsky's ATTRACT, for theta roles). The use of a combinator takes care of the syntax in parallel to the semantics, without the need for any other special devices. In addition, the operations are strictly local, so that some economy conditions are automatic.

Like these authors, we assume that obligatory control should be assimilated to raising, and we illustrate the combinator with a control structure.[15] We assume that **R** is an element of NL, like **A** and **B**. The examples shown give the LF ordering, which is assumed to be uniformly given by selection to the right, including the selection of DNPs for their predicates. The order prior to any other 'movement' then has DNPs preceding lexical heads, and all other complements following the head. We omit most functional heads, for expository purposes.

(27) [**A** Mary [**R** tried [to win]]]

Here, what we require is that the argument that eventually discharges the theta role of *try* is the same as that which discharges the theta role of *win*. We obtain this result by assigning the combinator the category, type and meaning given in (28).

(28) **R** category: X/D [/(Y/D) /(X/D/Y)] type: <(*a*/*b*)/(*c*/*b*)/(*a*/*b*/*c*)>

$(\mathbf{R}\, f\ \ g)\, x = (f\,.\,(g\,.\,x))\,.\,x$ where *f* has type <*a/b/c*> and *g* has type <*c/b*>

i.e. $\mathbf{R}\, f\, g = \lambda x\ [(f\,.\,(g\ .\,x))\ .\ x]$

Its occurrence is regulated by the syntactic Case system. We postulate that the head *likely* in (26) and the head *try* in (27) fail to case-license their propositional-type internal arguments as required for a CP argument. Following Cormack (1999), we assume that if and only if a lexical selection has a [+Case] feature may an argument be merged using the **A** combinator. If the selection has a [–Case] feature, then the partial argument discharge given by **R** is obligatory. If the selection has no Case feature, then no argument can be merged until a Case feature is supplied, for instance by finite T. The

[15] See also analyses in HPSG (Pollard & Sag section 3.5) and LFG (Bresnan 2000: chapter 12). Unlike our use of **R** here, these analyses have distinct levels at which the constraints giving the surface syntactic structure (using a VP) are stated, and that at which the connection between the external theta role of this VP and some other argument of local structure is stated.

Case licensing can be shown on the selection with a plus or minus sign.[16] The lexical entries for *try* and *likely* will be as in (29) and (30), where the primed forms are the extensionalised standard Montague-style denotations:

(29) *try* category: V [/D /V–]; semantic type: <*t*/*e*/*t*>; TRY = $\lambda p\ \lambda x\ [(\text{try}'.p).x)]$

(30) *likely* category: A [/D /V–]; semantic type: <*t*/nil/*t*>; LIKELY = $\lambda p\ \lambda x\ \text{likely}'.p$

The 'nil' role in (30) arises from the assumption that all external arguments have the category D, and that there is a 'nil' theta role assigned in the case of unaccusatives, including *likely*. This is perfectly respectable semantically: it is obtained from the standard meaning of *likely* via vacuous lambda abstraction, as shown. The meanings of *try to win* and *likely to fail* work out as shown in (31) and (32). The auxiliary *to* has identity semantics, and is ignored.

(31) [**R** TRY] TO-WIN = $\lambda \text{x}\ [(f.(g.x)).x]$

for f = TRY = $\lambda p\ \lambda y\ [(\text{try}'.p).\ y$ and g = TO-WIN = $\lambda z\ \text{win}'.z$

= $\lambda x\ [(\text{TRY}.\ (\lambda y\ \text{TO-WIN}.y).x)).x]$

= $\lambda x\ [(\lambda p\ \lambda y\ [(\text{try}'\ p).\ y)]\ .(\text{TO-WIN}\ .x)).x]$

= $\lambda x\ [(\lambda p\ \lambda y\ [(\text{try}'.p).\ y)]\ .(\ (\lambda z\ \text{win}'.z\ ).x)).x]$

= $\lambda x\ [(\lambda p\ \lambda y\ [(\text{try}'.p).\ y)]\ .(\text{win}'.x\ )).x]$

= $\lambda x\ [\lambda y\ [(\text{try}'.(win¢.\ x).y)].x]$

= $\lambda x\ [(\text{try}'\ win¢\ .x)\ .x]$

(32) [**R** LIKELY] TO-FAIL = $\lambda x\ [(f.(g.x)).x]$

for f = LIKELY = $\lambda p \lambda w\ \text{likely}'.p$ and g = TO-FAIL = $\lambda z\ \text{FAIL}.\ z$

= $\lambda x\ [(\text{LIKELY}\ .(\ \lambda z\ \text{FAIL}.z).x)).x]$

= $\lambda x\ [(\text{LIKELY}\ .(\text{FAIL}\ .x)).x]$

[16] For ease of reading, we have shown the Case feature on the c-selection category, although it would be more correct to show them on the types, since they affect meaning.

$$= \lambda x\ [(\lambda p \lambda w \text{ likely}'.p.(\text{FAIL } .x)).x]$$
$$= \lambda x\ [(\lambda w \text{ likely}'.(\text{FAIL } .x)).x]$$
$$= \lambda x\ [(\text{likely}'.(\text{FAIL } .x))]$$

The assimilation of control to raising means among other things that it is possible to conjoin control and non-control heads, as in (33):

(33) You [**can** and **will**] climb that hill

can: Aux[/D/V–], type <*t/e/t*> *will*: Aux[/D/V–], type <*t*/nil/*t*>

Other structures in which **R** is implicated include passive (where we need a trace), VP-anaphora, small clauses, object control and ECM (raising to object) structures, and *tough*-movement. Reconstruction can be handled by using the type lifting operator and higher-type traces. Simplified examples are given in (34) to (43), but we will not go through these (for more detail, see Cormack 1999). Informally, what R does is pass an external selection of some unsaturated internal argument up to the next selection. From here it will be discharged by the next argument supplied. For example in (36), the external role of *silly* will be discharged by the next selection by *find*. *Find* itself only requires this argument to saturate a nil role, so the argument *the play* saturates just the external role from *silly*, as required. With control examples such as (37), in contrast, the second argument, *Pip*, is induced by **R** to discharge both the external role from the VP *to leave*, and that supplied by *persuade*. In examples such as (39) and (40), the *Op* clause will require the **Gap** operator, and will be put together using 'composing' combinators such as **B**, which pass the internal role up the clause. *Op* itself is an identity operator, so it transmits this role, which now behaves like an external selection by the *for*-clause. For more detail on examples using higher type traces, see Cormack & Smith (2000b). What is important is that provided the heads are given the right selections, the merge induced by the combinator **R** does the right semantic and syntactic job, strictly locally.

(34) [**A** nothing [**R** was [**R** seen *t*]]]

(35) A: Mary tries to seem rich [[**A** MARY] [**R** TRY [**R** SEEM RICH]]]

B1: So does Bill [$_{VP}$ *e*] 'try to seem rich' : [**R** TRY [**R** SEEM RICH]]

B2: Alice actually does [$_{VP}$ *e*] 'seem rich' : [**R** SEEM RICH]

(36) [**A** Jim [**A** [the play] [**R** found [silly]]]] *find*: V[/D/D+/A–], type <*t*/*e*/nil/*t*>

(37) [**A** Grace [**A** Pip [**R** persuaded [to leave]]]] *persuade*: V[/D/D+/V–], <*t/e/e/t*>

(38) [**A** She [**A** [the tree] [**R** believe [to be old]]]] *believe*: V[/D/D+/V–], <*t*/e/nil/*t*>

(39) [**A** Jeremy [**R** is [**R** easy [*Op* [for anyone to deceive *t*]]]]]
easy: A[/D/C–], type <*t/*nil/*t*>

(40) [**A** The food [**R** is [**R** ready [*Op* [for us to eat *t*]]]]]
ready:A[/D/C–], type <*t/e/t*>

(41) [**A** [a unicorn] [**R** seems [**R** [to be - [eating the roses]]]]]]

reconstruction to the '-' position,
where - = $\lambda P \lambda \mathcal{P}\, \mathcal{P}.P$ for *P* of type <*t/e*> and $\mathcal{P}$ of type <*t*/(*t*/*e*)>.

(42) (in the context of an empty place at the table);

[**A** [someone] [**R** seems - [not to be have arrived]]]

(43) [A solution to the trisection problem] [was vainly [[- sought] *T*]]

reconstruction to the trace position,
where - = $\lambda P \lambda \mathcal{P}\, \mathcal{P}.P$ for *P* of type <*t/e/e*>and $\mathcal{P}$ of type <(t/e)/(*t/e/t*)>

7 Word order variation induced by R*

Suppose we assume, as we should on grounds of symmetry, that there is a composing version of **R**, which we label **R***.

(44) ((**R***. *f*).*g*).*x* = (**R**. *f*).(*g*.*x*) i.e. ((**R***. *f*).*g*) = λ*x* (**R**. *f*).(*g*.*x*)

In a purely head-initial language (i.e. one where all the arguments of a head appear to its right at LF), the use of a composing combinator, such as **B** rather than **A,** keeps the underlying word order constant, as we saw in example (5). According to the hypotheses about LF order we make, because binding noun-phrases are functors, selecting X/D to discharge a D selection, they appear before the relevant X/D phrase. The LF order arising will thus be 'SOVX', where X is the position of complements other than binding noun phrases. Under these circumstances, the use of a composing combinator induces differences in LF word order, as we see in the examples below. This means that certain word order variants which have been ascribed to movement can be explained rather by the choice of a composing combinator. We exemplify here in the simplest possible way using languages that are V-final at PF, so that the structure is not obscured by other displacement. We do not intend to commit ourselves to saying that no other 'displacement' is involved in such examples, nor that explanations or language particular restrictions are not needed to prevent overgeneration.

Consider first the examples in (45) from Nupe. The order in (45a) is the expected one, using **R**, where *kata tú* is the type <t/e> complement of the control verb *má* of type <t/e/e/t>, selecting initially for V[−Case]. If instead of **R**, we use **R***, then we can combine the control verb *má* directly with the verb *tú*, deferring the merge of the internal selection of *tú*. This will lead to the alternative (and preferred) verb-stacking Nupe form in (45b) (examples provided by Ahmadu Kawu).

(45) a [**A** musa [**B** á [**R** má [**A** kata tú]]]
Musa PERF know.how house build

b [**A** musa [**B** á [**A** kata [**R*** má tú]]]
Musa PERF house know.how build

'Musa has known how to build a house'

The Dutch example in (46) is comparable to the Nupe; but more spectacular examples such as (47), can be found. This can be analysed in the same way, with multiple use of **R***, as shown in (48).

(46) …dat zij appels moet eten

that she apples must eat

'… that she must eat apples'

(47) …dat ik Cecilia Henk de nijlparden zag helpen voeren
that I Cecilia Henk the hippopotamuses saw help feed

'… that I saw Cecilia help Henk feed the hippopotamuses'

(48) [**A** dat [**A** ik [**A** Cecilia [**A** Henk [**A** [de n.] [**R*** zag **R*** helpen voeren]]]]]]

zagen, *helpen*: V[/D/D+/V–], <*t*/*e*/*e*/*t*> (control structures);

voeren: V[/D/D+], <*t*/*e*/*e*>; [**R*** *helpen voeren*]: V[/D/D+/D+] , <*t*/*e*/*e*/*e*>;

[**R*** *zag* [**R*** *helpen voeren*]]: V[/D/D+/D+/D+], <*t*/*e*/*e*/*e*/*e*>

The examples above are from Steedman (2000: 136, 134), and the analysis follows Steedman in using composition. Steedman's chapter 6 is a detailed analysis of Dutch verb clusters in a standard Combinatory Categorial Grammar: it uses directional selection where our version allows Split Sign displacement, so that the required rules and restrictions will differ.

8 Discussion: Why do languages exhibit displacement?

With our analyses of the various forms of displacement sketched out, we are in a position to ask why natural languages do exhibit displacement. We begin with head-movement.

If we view "head-movement" as a reflex of Split Signs, we can see two possible reasons for its existence. The first is that it permits the expressive content of the various structures mentioned above, although to do this it must abandon the overt registration in syntax or morphology of scopal relations. Second, it minimises the number of words relative to the number of LF heads or meanings in a sentence or phrase (as does a lexical account of an inflected verb).[17] We see the motivation for this as external to the language faculty, and concerned primarily with the task of input processing. One of the first tasks in sentence processing is to ascertain word boundaries: complex lexical entries of this type then potentially reduce the complexity of this task. Morpheme boundaries too must be found: if the lexical entry is fusional rather than simply agglutinative, then the number of morpheme boundaries is reduced too.[18] Babies spend several months learning how to identify these boundaries (at least, from 7.5 to 10.5 months: Jusczyk 1999), so we assume that this represents a gain in on-line processing, but that such gains may come at the cost of an expanded number of lexical rules or lexical entries. There is indeed evidence that the input processing of regular inflectional morphology takes longer than the processing of irregular forms (McQueen & Cutler 1998:419; Ullman 2000: 147). That all processing of phonologically present material is costly relative to pragmatic processing is shown by the otherwise inexplicable existence in some languages of VP anaphora, or null subjects, and indeed by the frequent occurrence of phonologically null heads. Different languages may compromise differently in balancing the competing desiderata, even within a narrow field.

We offer some speculative reasons for the existence of the combinator **R** in Natural Language. We do not see the primary purpose as enabling the reorganisation of the natural topic:comment structure which occurs in passive and raising, which are often given as the *raison d'être* of A-movement. We surmise that the primary reason is connected with the 'control' instances, where what the combinator does is to form a complex predicate with at least one shared argument. The complex predicate is in the scope of a single finite T, and in most cases at least, forms a single event (though we have not shown how, here). Presumably, such complex predicates are within the range of the Language of Thought, perhaps expressed unitarily, or perhaps by means of **R**, or possibly by some other means not available in NL. There is an interesting question as to whether all such complex predicates would form legitimate lexical items, in the sense of

---

[17] Emonds (1994: 162) says that 'the most economical realisation of a given deep structure minimises the insertions of free morphemes'. Collins (2001: 58, 60) identifies such desiderata as economy principles relating to Spell Out, but there is a prerequisite that languages allow such forms in the lexicon and syntax. So far as the economy conditions are concerned, we take it that Jacobson (1998) is correct in arguing that these are due to processing and production principles, usually of the form of a 'race' to obtain a representation with the required properties.

[18] For instance, for PAST and SING, we assert that it is easier to process *sang* than *sing+ed*, and *sing+ed* than *past ... sing*.

having the right number of arguments with theta roles of possible kinds (e.g. not having two goals, or six arguments). Dutch examples such as that in (47) would clearly fail. We assume that restrictions on possible lexical entries aid acquisition. Either way, **R** must be an item of the internal language. These complex events are not talked of frequently enough, presumably, to justify independent lexicalisation in NL, even where possible, and in any case, would require complicated meaning postulates or inference rules. As is well-known, the head of the complex predicate does require its own Meaning Postulates — *persuade that* is not exactly the same *persuade* as *persuade to* — but at least most are the same, (and the new ones fall among a small number of options), while those of the second head are intact. We suppose then that the presence of **R** in NL is at least in part due to the demand for expressive power balanced against the demand for lexical restrictiveness.

Consider now A′ movement. We argued that this could occur essentially because of the availability of the combinator **B** (and, if **R** is used, **R*** will be needed too). It is reasonably clear that the availability of **B** in addition to **A** adds to the expressive power of a language. For example, as argued extensively by Steedman (1985, 2000), it enables us to conjoin non-standard constituents, as in (49), without any 'Right Node Raising' movement:

(49) a John likes and Mary dislikes no-one.

b [**A** NO-ONE [∧ [**B** JOHN LIKES] [*and* [**B** MARY DISLIKES]]]]

However, that is probably not the primary reason for the availability of **B**. Steedman argues that **B** is used to permit left to right parsing. This is almost certainly essential for the practical use of any NL which is head-initial, because of the processing disadvantage which would otherwise be involved in parsing predominantly right-branching structures. If **B** is available, then incremental left to right processing, which may including inference to enable on-line disambiguation, is possible. In the straightforward uses of this kind, as in example (50), exactly the same set of sentences are parsable as if we just had **A**, and nothing is added in the way of expressive power. However, once **B** is available, the way is open for A′-movement, as in (51), since the additional lexical items minimally required — **Gap**, and Traces — are within the range of what is already possible. However, as we have argued, the primary motivation for using **Gap**-fronted structures is to aid pragmatic processing.

(50) a [**A** Martha [**A** has [**A** often cried]] right-branching

b [**A** [**B** [**B** Martha has] often] cried] left-branching

(51) [**A** beans [**B** **Gap** [**B** she [**B** likes *t*]]]]

If the combinator **R** is available in NL, then the composing version, **R***, is to be expected, on the same grounds as we expect the composing version, **B**, of **A**. What is in need of explanation is not why it exists, but how it is exploited.

We have argued that none of the three classic cases of movement actually involve movement at all. Head-movement is replaced by the merger of Split Signs; A′-movement is replaced by merge of an NL trace in a lower position and the argument in the A-bar position; A-movement is replaced by the direct merger of the required constituents, mediated by a new combinator **R** instead of plain function argument application. We have suggested that the driving forces motivating these operations are of a variety of kinds, but we can see many of these as involving Interface requirements or desiderata (Chomsky 1995a, 2001), or other factors external to the linguistic system. However, the 'Interfaces' in question are not in general levels of representation, but processing devices (perhaps including transducers) forming links between on the one hand the LF and PF representations we are assuming, and on the other, the associated conceptual or perceptual/articulatory representations. For head-movement proper, we appealed to processing cost at the Perceptual Interface involved in identifying word and morpheme boundaries. For the phrasal version of head-movement, the proposal is that operations such as fronting are motivated by the need to reduce input processing costs at the Cognitive Interface, taking this to include language-directed pragmatic processing of linguistic input, used in obtaining the Explicature of the input (Carston 1999, 2000). For 'optional' A′-movement, we suggested a motivation again appealing to the Perceptual Interface, but this time turning on the temporal order in which linguistic material is presented to the hearer and the memory load otherwise accruing if the NL is head-initial. For A-movement, in contrast, we appealed in part to expressive power; that is, essentially to output demands at the Cognitive Interface. We also appealed to learnability: if the selection properties of lexical items are tightly constrained, then learning the syntactic and semantic properties of words is easier. This may be critical to childhood acquisition, when syntactic parameters as well must be set, and might be necessary for adults too. In addition, we suggested that some apparent displacement may be nothing but the LF ordering arising from the use of composing rather than direct combinators.

It is important to notice that what we have been suggesting as motivation above is motivation for the availability within UG of the enabling processes and items — Split Signs, F**on** and **Gap** heads, the composing combinators. Once these things are present in a particular language, they may be exploited for independent reasons. Most obviously, because F**on** is semantically null, it is exploited purely to obtain pragmatic effects; but such effects may be introduced for stylistic reasons, rather than to aid the retrieval of propositional contextual effects. Similarly, focus fronting in English is not obligatory, so the introduction of **Gap** for this purpose may have a stylistic rather than more directly communicative motivation. Even Split Signs may come to be exploited in such a fashion, where lexical entries permit. For example, the use of *will … not* rather than the easier *won't …* may be used to signal a degree of formality.

There is no one perfect solution for a NL to meet the legibility conditions imposed by the interfaces. Rather, the interaction of various often incompatible desiderata gives rise to a number of possible compromise solutions. Whether any particular solutions within this conceptual space are optimal is an open question.

In conclusion we hope we have shown that we should take seriously the idea that the philosophy and findings of P&P grammars and the Minimalist program can be economically instantiated without recourse in the grammar to actual movement, nor to its pseudo-instantiation in Copy Theory or Internal Merge. We have argued that the additional apparatus necessary to add the apparent displacements to the grammar has external motivation.

References

Abasheikh, M. (1978) *The Grammar of Chimwi:ni Causatives*. PhD, University of Illinois.

Ackema, P. & A. Neeleman (2000a) On the different selection properties of null and overt affixes. *GLOW Newsletter* 44: 62-63.

Ackema, P. & A. Neeleman (2000b) M-selection and phrasal affixation.. In *UCL Working Papers in Linguistics* 12: 307-342.

Alsina, A. (1999) Where's the mirror principle? *The Linguistic Review* 16: 1-42.

Baker, M. (1985) The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16: 373-415.

Bresnan J. (2000) *Lexical-Functional Syntax*. Oxford: Blackwell.

Brody, M. (1990) Some remarks on the focus field in Hungarian. In *UCL Working Papers in Linguistics* 2: 201-225.

Carston, R. (1999) The relationship between generative grammar and (relevance-theoretic) pragmatics. *UCL Working Papers in Linguistics* 11:21-39.

Carston, R. (2000) Explicature and semantics. *UCL Working Papers in Linguistics* 12:1-44.

Chomsky, N. (1977) On wh-movement. In Culicover, P.W., T. Wasow & A. Akmajian (eds.). *Formal Syntax*. 71-132. New York: Academic Press.

Chomsky, N. (1995a) *The Minimalist Program.* Cambridge MA: MIT Press.

Chomsky N. (1995b) Bare Phrase Structure. In Webelhuth, G. (ed.). *Government and Binding Theory and the Minimalist Program*. 383-439. Oxford: Blackwell.

Chomsky, N. (1999) Derivation by phase. ms., MIT.

Chomsky, N. (2000). Minimalist inquiries: the framework. In Martin, R., D. Michaels & J. Uriagereka (eds.). *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik* . 89-155. Cambridge MA: MIT Press.

Chomsky, N. (2001) Beyond explanatory adequacy. Ms., MIT.

Cinque, G. (1977) The movement nature of left dislocation. *Linguistic Inquiry* 8: 397-412.

Cinque, G. (1990) *Types of A-bar dependencies*. Cambridge MA: MIT Press.

Collins C (2001) Economy conditions in syntax.. In Baltin, M. & C. Collins (eds.). *The Handbook of Contemporary Syntactic Theory*. 45-61. Oxford: Blackwell.

Cormack A. (1982) VP Anaphora: variables and scope. In Landman, F. & F. Veltman (eds.). V*arieties of Formal Semantics*. 81-102. Dordrecht: Foris.

Cormack, A. (1989) *The Syntax and Semantics of Definitions*. PhD, University of London. Published (1998) as *Definitions: Implications for Syntax, Semantics, and the Language of Thought*. New York: Garland.

Cormack, A. (1999). Without Specifiers. In Adger, D., S. Pintzuk, B. Plunkett & G. Tsoulas (eds.). *Specifiers: Minimalist Approaches*. 46-68. Oxford: Oxford University Press.

Cormack, A. & N.V. Smith (1999) Where is a sign merged? *Glot International* 4: 20.

Cormack, A & N.V. Smith (2000a) Head movement and negation in English. *Transactions of the Philological Society* 98: 49-85

Cormack, A. & N.V. Smith (2000b) Fronting: the syntax and pragmatics of 'Focus' and 'Topic'. *UCL Working Papers in Linguistics* 12: 387-416.

Cresti, D. (1995) Extraction and reconstruction. *Natural Language Semantics* 3: 79-122.

Dalrymple, M., R.M. Kaplan, J.T. Maxwell & A. Zaenen (1995). *Formal Issues in Lexical-Functional Grammar*. Stanford: CSLI.

Emonds J. (1994) Two principles of economy. In Cinque, G., J.-Y. Pollock, L. Rizzi, & R. Zanuttini (eds.). *Paths towards UG: studies in honor of Richard Kayne*. 155-172. Washington DC: Georgetown University Press.

Fodor, J.A. (1975) *The Language of Thought*. Harvard University Press.

Fox, D. (1999) Reconstruction, Binding Theory, and the interpretation of chains. *Linguistic Inquiry* 30: 157-196.

Frazier, L. (1988) Grammar and language processing. In Newmeyer, F. (ed.). *Linguistics: The Cambridge Survey* Vol II. 15-34. Cambridge: Cambridge University Press.

Grimshaw, J. (1997) Projection, heads, and optimality. *Linguistic Inquiry* 28: 373-422.

Heim, I. & A. Kratzer (1998) *Semantics in generative grammar*. Oxford: Basil Blackwell.

Heycock, C. & A. Kroch (1999) Pseudocleft connectedness: implications for the LF interface. *Linguistic Inquiry* 30: 365-397.

Hornstein, N. (1998) Movement and chains. *Syntax* 1: 99-127.

Hornstein, N. (1999) Movement and control. *Linguistic Inquiry* 30: 69-96.

Hornstein, N. (2001) *Move! A Minimalist Theory of Construal.* Oxford: Basil Blackwell.

Hyman, L.M. (2001) Templatic and/or compositional morphology: suffix ordering in Bantu. Handout from lecture given at UCL.

Hyman, L.M. & Mchombo, S. (1992) Morphotactic constraints in the Chichewa verb stem. *Berkeley Linguistics Society* 18: 350-364.

Jackendoff, R. (1972) *Semantic Interpretation in Generative Grammar*. Cambridge MA: MIT Press.

Jackendoff, R. (1997) *The Architecture of the Language Faculty*. Cambridge MA: MIT Press.

Jacobson, P. (1990) Raising as Function Composition. *Linguistics and Philosophy* 13: 423-476.

Jacobson, P. (1992) Raising without movement. In Larson, R., S. Iatridou, U. Lahiri & J. Higginbotham (eds.). *Control and Grammar*. Dordrecht: Kluwer.

Jacobson, P. (1998) Where (if anywhere) is transderivationality located? In Culicover, P.W. & L. McNally (eds.). *Syntax and Semantics 29: The Limits of Syntax*. 303-336. San Diego: Academic Press.

Jacobson, P. (1999) Towards a Variable-Free Semantics. *Linguistics and Philosophy* 22: 117-184.

Jusczyk, P.W. (1999) How children begin to extract words from speech. *Trends in Cognitive Science* 3: 323-328.

Kayne, R.S. (1994) *The Antisymmetry of Syntax.* Cambridge MA: MIT Press.

Krifka, M. (1991) A compositional semantics for multiple focus constructions. *SALT I*: 127-158.

Ladusaw, W.A. (1980) *Polarity Sensitivity as Inherent Scope Relations*. New York: Garland.

Lechner, W. (1998) Two kinds of reconstruction. *Studia Linguistica* 52: 276-310.

Manzini, M.R. & A. Roussou (2000) A minimalist theory of A-movement and control. *Lingua* 110: 409-447.

McQueen, J.M. & Cutler, A. (1998) Morphology in word recognition. In Spencer, A. & A.M. Zwicky (eds.). *The Handbook of Morphology.* 406-427. Oxford: Blackwell.

Muysken, P. (1981) Quechua word order. In Heny, F. (ed.). *Binding and Filtering.* 279-327. Cambridge MA: MIT Press.

Pierrehumbert, J. (1980) *The Phonology and Phonetics of English Intonation*. PhD, MIT.

Pollard, C. & I. A. Sag (1994) *Head-driven Phrase Structure Grammar*. Chicago: CSLI, University of Chicago Press.

Pullum, G.K. (1976) The Duke of York gambit. *Journal of Linguistics* 12: 83-102.

Rizzi L. (1997) The fine structure of the left periphery. In Haegeman, L. (ed.). *Elements of Grammar: Handbook in Generative Syntax*. 281-337. Dordrecht: Kluwer.

Roberts I. (1998) Have/be Raising, Move F, and Procrastinate. *Linguistic Inquiry* 29, 113-125.

Romero, M. (1998) The correlation between scope reconstruction and connectivity effects. *Proceedings of WCCFL XVI.*

Rooth, M. (1992) A theory of focus interpretation. *Natural Language Semantics* 1: 75-116.

Sauerland, U. & P. Elborne (1999). Total reconstruction, PF-movement, and derivational order (draft). ms. Tübingen and MIT.

Sperber, D. & D. Wilson (1995) *Relevance: Communication and Cognition* (second ed.). Oxford: Basil Blackwell.

Steedman, M. (1985) Dependency and Coordination in the grammar of Dutch and English. *Language* 61: 523-568.

Steedman, M. (1993) Categorial Grammar. *Lingua* 90: 221-258.

Steedman, M. (1996) *Surface Structure and Interpretation*. Cambridge MA: MIT press.

Steedman, M. (2000) *The Syntactic Process*. Cambridge MA: MIT press.

Stump, G.T. (1996) Template morphology and inflectional morphology. In Booij, G. & J. van Marle (eds.). *Yearbook of Morphology*. 217-241. Dordrecht, Kluwer.

Szabolcsi, A. (1992) Combinatory Grammar and projection from the lexicon. In Sag, I. & A. Szabolcsi (eds.). *Lexical Matters*. 241-268. Chicago: CSLI, University of Chicago Press.

Ullman, M. (2000) A mental model of morphology: The psychological and neural bases of the computation of complex words. *University of Maryland Working Papers in Linguistics* 10: 127-156.

Zanuttini R. (1997) *Negation and Clausal Structure: a Comparative Study of Romance Languages*. Oxford: Oxford University Press.